



Evaluation of sgRNA Target Sites for CRISPR-Mediated Repression of *TP53*

Ingrid E. B. Lawhorn, Joshua P. Ferreira, Clifford L. Wang*

Department of Chemical Engineering, Stanford University, Stanford, California, United States of America

Abstract

The CRISPR (clustered regularly interspaced short palindromic repeats) platform has been developed as a general method to direct proteins of interest to gene targets. While the native CRISPR system delivers a nuclease that cleaves and potentially mutates target genes, researchers have recently employed catalytically inactive CRISPR-associated 9 nuclease (dCas9) in order to target and repress genes without DNA cleavage or mutagenesis. With the intent of improving repression efficiency in mammalian cells, researchers have also fused dCas9 with a KRAB repressor domain. Here, we evaluated different genomic sgRNA targeting sites for repression of *TP53*. The sites spanned a 200-kb distance, which included the promoter, transcript sequence, and regions flanking the endogenous human *TP53* gene. We showed that repression up to 86% can be achieved with dCas9 alone (i.e., without use of the KRAB domain) by targeting the complex to sites near the *TP53* transcriptional start site. This work demonstrates that efficient transcriptional repression of endogenous human genes can be achieved by the targeted delivery of dCas9. Yet, the efficiency of repression strongly depends on the choice of the sgRNA target site.

Citation: Lawhorn IEB, Ferreira JP, Wang CL (2014) Evaluation of sgRNA Target Sites for CRISPR-Mediated Repression of *TP53*. PLoS ONE 9(11): e113232. doi:10.1371/journal.pone.0113232

Editor: Robertus A. M. de Bruin, University College London, United Kingdom

Received: September 8, 2014; **Accepted:** October 21, 2014; **Published:** November 14, 2014

Copyright: © 2014 Lawhorn et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. All qRT-PCR and flow cytometry data are available from the Stanford Digital Repository (<http://purl.stanford.edu/rq049bp8124>).

Funding: IEBL was funded by the National Science Foundation Graduate Research Fellowship Program <http://www.nsfgrfp.org/> and Stanford Graduate Fellowship <http://sgf.stanford.edu/>. IEBL, JPF, and CLW were funded by the National Institute on Aging (NIH-5R21AG040360-02) <http://www.nia.nih.gov/>. Data and sorting was performed on instruments in the Stanford Shared FACS Facility obtained using National Institutes of Health S10 Shared Instrumentation Grants S10RR025518-01 and S10RR027431-01 http://dpcpsi.nih.gov/orip/diic/shared_instrumentation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* Email: cliff.wang@stanford.edu

Introduction

CRISPR (clustered regularly interspaced short palindromic repeats) serves as an adaptive immune system for many bacteria and archaea [1,2]. In the *Streptococcus pyogenes* type II CRISPR system, CRISPR-associated 9 nuclease (Cas9) binds to an RNA complex of trans-acting RNA (tracrRNA) and CRISPR RNA (crRNA), which guides the complex to DNA sequences complementary to the crRNA. This Cas9-RNA complex can then recognize and cleave the DNA of pathogens or other foreign elements [3–5]. To develop a general, host-independent method for targeting Cas9 and other recombinant proteins to desired DNA sequences, researchers have devised a platform that utilizes a small guide RNA (sgRNA) consisting of the two RNA elements joined by a short hairpin [5]. Targeting is achieved by the base-pairing between a sequence element encoded by the sgRNA and a desired, target DNA sequence. Any sequence can be targeted as long as it lies upstream of a protospacer adjacent motif (PAM) consisting of the sequence NGG (N represents any nucleotide base) [6,7]. Thus far, researchers have employed this technology to achieve targeted genomic insertions, deletions, and mutations in bacteria and eukaryotic cells [8–14].

More recently, researchers have adapted the CRISPR technology to deliver proteins of interest to targeted DNA sequences for applications beyond genome editing. By inactivating the two endonucleolytic domains of Cas9 to create a catalytically “dead”

Cas9 (dCas9) protein [5], researchers have employed the CRISPR gene-targeting platform and conjugated effector domains to achieve transcriptional repression [15–18], activation [18–21], and genomic loci imaging [22]. The CRISPR-mediated repression of gene expression has been termed CRISPR interference (CRISPRi). With CRISPRi, when dCas9 was co-expressed with sgRNA that was complementary to targeted gene, up to 99.7% transcriptional repression was observed in bacteria [15]. Such repression was achieved by targeting the non-template DNA strand of the gene coding region to block or hinder RNA polymerase transcription elongation [23]. Alternatively, repression has been achieved by targeting sites within a gene promoter region to block RNA polymerase and transcription factor binding for transcription initiation.

In contrast to repression of genes in prokaryotes, recent reports of dCas9-mediated repression in mammalian cells have been more modest, with repression reported to be approximately 50%. To improve on the repression reported in mammalian cells, Gilbert *et al.* fused a Krüppel-associated box (KRAB) transcription repressor domain from Kox1 to the C-terminus of dCas9 and multiple sgRNA target sites against the same gene locus were evaluated [16]. While the KRAB effector domain generated up to 93% repression of a reporter gene expression in human cells, the extent of endogenous gene knockdown varied from gene to gene with, in the best case, up to 80% repression of transferrin receptor CD71 at the protein level [16].

The rules for achieving CRISPR-mediated repression in mammalian systems are still not clearly established. When targeting the transcribed regions of a gene, in bacteria it has been reported that targeting the non-template strand leads to greater repression than when targeting the template strand. Based on this finding, it has been previously recommended that only the non-template strand should be targeted for repression in mammalian cells [23], though to our knowledge this has not been explicitly evaluated outside of bacterial systems. There is evidence that the extent of repression can be greater when targeting DNA sequences close to the transcriptional start site (TSS) in bacteria [15] and human embryonic stem cells [18] than sites in the transcribed region. However, to our knowledge a systematic comparison of promoter-proximal and transcribed region sites has not yet been conducted in mammalian cells. Last, while the KRAB effector domain has been shown to be more effective than dCas9 alone in a select number of genes [16], it is unclear whether the use of the KRAB effector domain is required to achieve strong endogenous repression.

A better understanding of how the choice of dCas9 target sites affects repression of mammalian gene expression would help those who seek to control gene expression for research or synthetic applications [24]. To address these issues, we evaluated the effect of dCas9-mediated transcriptional repression by using sgRNAs complementary to a variety of target sites covering a 200 kb distance: these sites include the promoter, transcribed sequence and regions adjacent to the endogenous *TP53* gene locus on both the template and non-template DNA strands. p53, the gene product of *TP53*, is a transcription factor whose activation by DNA damage or other cellular stresses results in responses that promote cell cycle arrest, apoptosis, or other tumor suppressing processes [25] and has, previously and presently, served as a well-studied model gene for the purpose of studying gene repression [26,27]. We also evaluated long-distance repression achieved by dCas9-KRAB binding in regions flanking the *TP53* gene locus. Finally, we evaluated whether repression could be achieved by targeting sites surrounding the transcriptional start site in other genes.

Materials and Methods

Vector design

Target sequences were designed using the protocol recommended in Larson *et al.* [23]. Briefly, targets of 20–25 nucleotides in length preceding an NGG PAM site were screened for off-target homology using NCBI BLAST and the CRISPR design tool on crispr.mit.edu. Targets were designed to bind to areas of interest within the human *TP53* promoter as well as to both the template and non-template strands of *TP53* exons 4, 7, and 10. Because the U6 promoter was used for sgRNA expression, all target sequences without a “G” nucleotide in the leading position had one appended to the 5′ end of the target sequence to facilitate expression. Additionally, sgRNA sequences were designed to target sites approximately every 10 kilobases away from the transcriptional start site of *TP53* in both the upstream and downstream directions. GFP targets against the non-template strand were created based on the sequences published in Gilbert *et al.* (GFP1) [16] and Mali *et al.* (GFP2) [10]. Additional targets were designed to bind to locations 20 to 50 base pairs upstream or 100–120 base pairs downstream of the transcriptional start site (+1) as identified by the DataBase of Transcriptional Start Sites [28,29] for various genes of interest in HEK 293 cells (Sequence Read Archive Accession No. SRA003625) [30]. Scramble1 and Scramble2 “target” controls were designed to have low homology

to the human or murine genome and are used as a non-targeting control to normalize gene expression. The sgRNA sequences are summarized in Tables S1, S2, and S3.

Annealed oligonucleotides containing the sgRNA sequences were ligated into the plasmid pRSET B-U6 sgRNA-term that was digested with BpiI (ThermoScientific, Pittsburgh, PA, USA). This step results in a RNA polymerase III U6 promoter-driven sgRNA expression cassette consisting of a 20–25 nucleotide domain complementary to a target DNA region, a Cas9-binding hairpin, and a transcription terminator. The U6-sgRNA expression construct was then digested with PspOMI and NotI-HF (New England BioLabs, Ipswich, MA, USA) and ligated into pCru5-1.2.1-BFP-IRES-Blast that was digested with NotI-HF to create pCru5-BFP-IRES-Blast-U6 sgRNA. Mammalian codon-optimized *Streptococcus pyogenes* dCas9 sequence with nuclease-inactivating D10A and H840A substitutions [5] (Addgene plasmid 44246) was fused to three C-terminal NLS (DPKKKRV) sequences alone or with the KRAB domain sequence. The entire construct was then inserted into pCru5-1.2-mEGFP-IRES-mCherry-F2A-Puro [31] digested with SphI-HF and NsiI (New England BioLabs) to form pCru5-dCas9-3xNLS-IRES-mCherry-F2A-Puro or its variant with the KRAB domain.

Cell culture and transfection

HEK 293 [32] and HEK 293T [33] cells were cultured in DMEM with 10% FBS. Media was supplemented with 1 mM glutamine, 100 U/mL penicillin, and 100 µg/mL streptomycin.

HEK 293 and HEK 293T cells were transfected using polyethylenimine (PEI) transfection. Cells plated in 6-well plates were transfected with 2 µg of the dCas9 expression plasmid and 1.2 µg of the sgRNA expression plasmid per well which can be scaled by surface area for larger or smaller plates. Cells transfected with pSuper or pSuper-p53 plasmids were co-transfected with 0.5 µg pCru5-1.2-GFP-IRES-mCherry-F2A-Puro or pEGFP-N3 for selection or sorting purposes. Multiplexed transfection mixes contained equal amounts of all sgRNA expression plasmids used. 1 day post-transfection, cells were split 50% and selected with 2 µg/mL puromycin.

HEK 293 cells were transduced with retrovirus for stable expression of destabilized GFP. Retrovirus was produced by co-transfecting a pCru5 plasmid and pCL-Ampho into HEK 293T cells using calcium phosphate precipitation. One pCru5 plasmid had a long terminal repeat (LTR) containing a promoter that drives transcription of the neomycin resistance gene and an additional mutated EF-1 α promoter (EF-T05) [34] driving expression of destabilized monomeric green fluorescent protein (GFP). The other pCru5 plasmid expressed GFP followed by an internal ribosome entry site (IRES) and a neomycin resistance gene directly off the LTR with a translation initiation sequence variant (1.2) [31]. Virus-containing supernatant was harvested and used to transduce HEK 293 cells. Virus was titered so that transduced cells received a single copy of the vectors. Polybrene (hexadimethrine bromide) was added to cultures at a concentration of 8 µg/mL. 24–48 h post-infection, cells were selected and maintained with 400 µg/mL neomycin.

Flow cytometry

Three days after antibiotic selection, cells were trypsinized, washed once with cold PBS, and resuspended in 1% FBS in PBS. For measurement of GFP knockdown, cells were analyzed on a LSRII (BD Biosciences, San Jose, CA, USA) and gated for mCherry- and TagBFP-positive viable singlet cells. GFP fluorescence on the FITC-A filter was recorded and GFP expression relative to the Scramble1 control was analyzed using FlowJo

software (Tree Star, Ashland, OR, USA). For indicated qRT-PCR experiments, samples were sorted on a FACSAria II (BD Biosciences) for TagBFP-positive or GFP-positive viable singlet cells on the Pacific Blue-A filter using the 85 μ M nozzle. At least 400,000 cells were sorted per sample.

Quantitative RT-PCR

Cell samples were trypsinized, washed once with cold PBS, and resuspended in RLT buffer supplemented with β -mercaptoethanol before freezing at -80°C for up to a week. Total RNA was isolated using the RNeasy Mini Kit (Qiagen, Valencia, CA, USA) using the manufacturer's directions. RNA was converted to cDNA using the High-Capacity cDNA Reverse Transcription Kit with random primers using the manufacturer's directions (Applied Biosystems, Foster City, CA, USA). All quantitative PCR reactions were done at 15 μ L volume using SYBR Green PCR Master Mix (Applied Biosystems). Each independent biological replicate was plated in triplicate and run on the StepOnePlus thermal cycler (Applied Biosystems). Table S4 lists the primer sequences used to measure expression of *TP53* (GenBank Accession No. NM_000546), *WRAP53 α* isoform (GenBank Accession No. NM_001143991), *WRAP53* all isoforms (GenBank Accession No. NM_018081), *GAPDH* (GenBank Accession No. NM_001256799), *MAPK1* (GenBank Accession No. NM_138957), *MAPK14* (GenBank Accession No. NM_001315), *PPIB* (GenBank Accession No. NM_000942), and *RBI* (GenBank Accession No. NM_000321).

Expression levels of housekeeping gene *TIMM17B* (GenBank Accession No. NM_001167947) were used for normalization of human gene expression. Relative quantity was calculated by the comparative C_T method ($\Delta\Delta C_T$) using cells transfected with the dCas9 construct and a Scramble sgRNA construct as the control. All reported levels of repression are relative to a Scramble control. Standardization between three independent biological replicates was performed via log transformation, mean centering, and autoscaling described by Willems *et al.* [35]. Relative expression levels from sorted experiments were correlated to their respective target's nucleotide annealing temperature using a linear regression for the target sequence (Figure S4).

Immunoblotting

To quantify p53 abundance, immunoblotting was performed using standard protocols. Endogenous p53 level was detected in lysate of transfected HEK 293T cells using a mouse anti-p53 primary antibody (1C12, #2524, Cell Signaling Technology, Danvers, MA, USA) and a goat anti-mouse IgG antibody conjugated to horseradish peroxidase (HRP). Loading control GAPDH was detected using a rabbit primary antibody conjugated to HRP (14C10, #3683 Cell Signaling Technology, Danvers, MA, USA). HRP activity was measured with the ECL Plus Western Blotting Detection Kit (#RPN2132, GE Healthcare, Piscataway, NJ, USA).

Results

Our goal was to investigate how the location of the sgRNA target site within a gene locus affects dCas9-mediated repression in human cells. We constructed a vector that expressed a human codon-optimized dCas9 protein with three nuclear localization signals (3xNLS). The vector also utilized an internal ribosome entry site (IRES) that allowed bicistronic expression of the dCas9 with mCherry-2A-Puro, which encoded both the mCherry red fluorescent protein and a puromycin resistance gene (Figure S1A). To evaluate the efficiency of the KRAB effector domain, a version

was also created with KRAB fused to the C terminus of the dCas9-3xNLS fusion (Figure S1A). Each sgRNA was expressed from a RNA polymerase III U6 promoter downstream of a cassette expressing blue fluorescent protein (BFP) and a blasticidin resistance gene (Figure S1A). To test the efficacy of our dCas9 fusions, we co-transfected the dCas9 constructs and individual sgRNAs in HEK 293 cells stably expressing green fluorescent protein (GFP). We observed reduction of GFP fluorescence over the non-targeted Scramble1 control in cells expressing a sgRNA targeting GFP and either dCas9 (37% repression) or dCas9-KRAB (60% repression) (Figure S2). As observed previously [16], dCas9-KRAB generated greater GFP repression than dCas9 alone.

In this study, we were interested in evaluating the repression of endogenous genes. Having established that our constructs could repress ectopic GFP expression, we next evaluated the effectiveness of dCas9 and dCas9-KRAB in repressing human *TP53* mRNA expression. To evaluate the effect of target site position on repression, we tested sgRNAs that targeted different sites at or near the *TP53* gene and measured *TP53* mRNA levels by quantitative real-time PCR (qRT-PCR). We hypothesized that the KRAB fusion may have long-distance repressive effects on gene expression since KRAB-mediated promoter silencing has been reported over distances of several tens of kilobases away from the KRAB binding site [36]. To evaluate this hypothesis, we designed sgRNAs to target sites up to 100,000 bp upstream and downstream of the *TP53* transcriptional start site (TSS) in approximately 10,000 bp increments (Figure S1B). Each sgRNA is labeled with a number representing the distance (bp) from the transcriptional start site ("K", one thousand bp) and a "-" or "+" representing upstream or downstream of +1. HEK 293T cells were co-transfected with a dCas9 or dCas9-KRAB expression construct and each sgRNA expression construct, selected with puromycin, and sorted for BFP positive cells. Of the 21 sgRNAs tested, the only appreciable repression of mRNA occurred at a site approximately 36 bp upstream from the *TP53* TSS (85% for dCas9; 90% for dCas9-KRAB; -T36, Figure S1C and Figure S1D). Furthermore, there was little observable additional *TP53* repression with the dCas9-KRAB fusion in comparison to repression with dCas9. Since using multiple sgRNAs has been successful in CRISPR-directed gene activation [19,20], we tested the effect of using multiple sgRNAs for repression by co-transfecting dCas9 or dCas9-KRAB with various combinations of sgRNAs targeting: 10 sites spanning 100,000 bp upstream of *TP53* (-T100K to -T10K), three sites spanning 30,000 bp upstream of *TP53* (-T30K to -T10K), 10 sites spanning 100,000 bp downstream of *TP53* (+T10K to +T100K), three sites spanning 30,000 bp downstream of *TP53* (+T10K to +T30K), and three sites flanking the *TP53* TSS (-T36 +T27 +T110). Again, there was no added transcriptional repression with dCas9-KRAB directed binding as compared to dCas9 (Figure S1E). Moreover, we observed similar repression of *TP53* with the simultaneous targeting of three sites (-T36 +T27 +T100) as with the targeting of the single -T36 site (Figure S1D). This is in contrast to the previously-observed synergistic increase in gene activation induced by use of multiple sgRNA binding sites in conjunction with a dCas9 fused to aVP64 transcription activation domain [20], tetramer of VP16 domains, each capable of recruiting transcriptional machinery to promoter proximal regions. We postulate that this synergy in activation occurred because activation by VP64 can act over long distances [37] thus allowing multiple targeted sites to act together from various distances. In contrast, the repression that we observed may be due to a TSS-proximal mechanism, where additional dCas9 repression factors cannot cooperate significantly to improve the degree of

repression. Our results also suggest that fusing KRAB to dCas9 does not significantly improve *TP53* transcriptional repression and that targeting multiple sites concurrently does not necessarily improve the efficiency of repression. However, we cannot rule out the possibility that KRAB may increase repression when targeted to other genes or using alternative sites.

Since dCas9-KRAB had little added benefit over dCas9 for repression of *TP53* and there was significant repression when targeting sites near the TSS, we further explored the positional effect of targeting dCas9 to sites closer to the *TP53* gene itself. 21 sgRNAs were designed to target sequences on the template and non-template strands of the promoter and transcribed region of *TP53* (Table S1 and Figure 1A and B). In bacteria, only sgRNAs binding to the non-template strand could enable repression of a fluorescent protein [15] although that finding has not been replicated in mammalian cells. The sgRNAs were co-expressed with dCas9 in transfected and puromycin-selected HEK 293T cells and directly analyzed by qRT-PCR. As before, the -T36 target site was shown to significantly repress *TP53* expression at the mRNA level ($P < 0.01$, Figure 1C) as compared to cells expressing either non-targeted control. Because transfection efficiency varied, we opted to sort for BFP-positive cells after transfection and puromycin selection to enrich for cells expressing both CRISPR constructs. Of the subset of sgRNAs tested, three repressed *TP53* transcriptional expression (Figure 1D), all located within a 200-bp region of the TSS. sgRNA -T36 induced approximately 86% *TP53* mRNA repression, a level comparable to shRNA knock-down with pSuper-p53 construct [26] (Figure S3C). Repression using this target site was also consistently observed at the protein level in unsorted, selected cells (Figure 1E) for three independent experiments. However, when targeting the +T110 and +T111 sites and analyzing protein levels by immunoblotting, not only were the magnitudes of repression lower than -T36 but they were also less reproducible over those same experiments. Additionally, the targeting of sites +T27 and +T110 repressed *TP53* transcription expression. We did not observe significant repression when targeting the template or non-template strands of DNA in the transcribed region. It should be noted that first exon of *TP53* overlaps with exon 1 α of *WRAP53* [38,39] which is the first exon of the *WRAP53* α isoform—a known anti-sense transcript of *TP53* (Figure S3A). We observed repression of the *WRAP53* α isoform in cells transfected with dCas9 and the -T36 sgRNA but no similar repression of total *WRAP53* expression from all three isoforms (Figure S3B); we did not see a similar effect with pSuper-p53 shRNA transfection (Figure S3C). Finally, we note that for the many sgRNAs in our study that did not generate appreciable repression of *TP53*, we cannot rule out that these sgRNAs were somehow incapable of mediating the binding of dCas9 to the DNA.

Observing that target sites -T36 and +T110 were effective at repressing *TP53* transcription, we next designed sgRNAs to target sites at similar distances from the TSS for five additional human genes: *GAPDH*, *PPIB*, *MAPK1*, *MAPK14*, and *RBI* (Figure 2). For each gene, dCas9 was targeted to one site 27 to 52 bp upstream of the TSS on the non-template DNA strand and two sites 81 to 136 bp downstream of TSS on the both the non-template and template DNA strands (Table S3). While it is certainly possible that the KRAB fusion could be effective for repression of these other genes, because we found the dCas9 alone was sufficient for repression of *TP53*, we chose to investigate the whether dCas9 alone could also be effective in repressing other genes. For four of the five genes, modest transcriptional repression was achieved, including over 50% repression in *RBI* with target +R121 (Figure 2F). To explain the variation in repression efficiency,

we hypothesized that the annealing temperature of the target base pair sequence may correlate with the efficiency of repression. Analysis of this parameter and the relative expressions for each target examined in triplicate revealed no correlation between the sgRNA target annealing temperature and mRNA expression (Figure S4). These results indicate modest transcriptional repression can be achieved in a variety of genes by targeting the region proximal to the TSS although there may be additional factors limiting the extent of that knockdown.

Discussion

Prior reports [16,23] recommended targeting either the promoter-proximal region or the transcribed region on the non-template strand for optimal endogenous dCas9 repression in human cells. In the case of *TP53*, while we were unable to achieve significant repression by targeting various sites on either the template or non-template strand of the transcribed region, we were able to repress transcription by targeting sites near or close to the TSS. We report up to 86% repression of *TP53* transcription by using sgRNAs -T36 and +T110, targeting sites 36 bp upstream and 110 bp downstream of the transcriptional start site. Interestingly, the -T36 sgRNA binds to a site that overlaps with previously-reported Myc/Max, USF, and E2F1 binding motifs while +T110 binds to a site overlapping with a Pax binding motif [40]. The -T36 sgRNA is also within 40 bp of the transcription start site, which is the approximate DNA footprint of RNA polymerase II [41]. In support of the notion the dCas9 can interfere with transcription at the DNA level, recent chromatin immunoprecipitation (ChIP) data by others indicate that sgRNA-mediated dCas9 binding at the transcriptional enhancer of *Nanog* interferes with the binding of several transcription factors in mouse ESCs [42]. The CRISPR-mediated repression that we observed is similar to what we were able to achieve using a *TP53* shRNA construct [26] and similar to levels of repression Gilbert *et al.* achieved by targeting sites downstream of the TSS of endogenous *CD71* and *CXCR4* genes [16]. We did not find a clear association between target proximity to the TSS and transcriptional repression, but we were able to achieve modest but significant repression in an additional four of five examined genes.

Interference with transcription initiation provides an alternative to post-transcriptional regulation via RNAi and allows for evaluation of repression at the transcriptional-level. In contrast to gene editing with RNA-directed Cas9, dCas9-mediated repression allows for a reduction in gene expression without modification to the genomic sequence. Thus with dCas9-mediated repression, it should be possible to engineer conditional expression system where repression can be switched “on” or “off” reversibly by other cellular cues or experimental conditions (e.g., addition of a small molecule inducer). In addition, it appears that, depending on the targeted gene, similar levels of transcriptional repression can be achieved without the use of a KRAB effector domain, which, for *TP53* transcription, appeared to have little effect. It is important to note potential limitations in the dCas9-initiated repression. Access of the dCas9-RNA complexes to the target site may be limited or even obstructed by the structure and local environment of the chromatin, epigenetic features, or the presence of regulatory elements [43,44] and thus affect efficiency of repression. Additionally, even if a sequence has few if any off-target binding sites [44–48], some promoter regions may overlap with other genes. We observed concomitant repression of *WRAP53* α isoform mRNA levels (Figure S3B) when targeting *TP53*'s TSS.

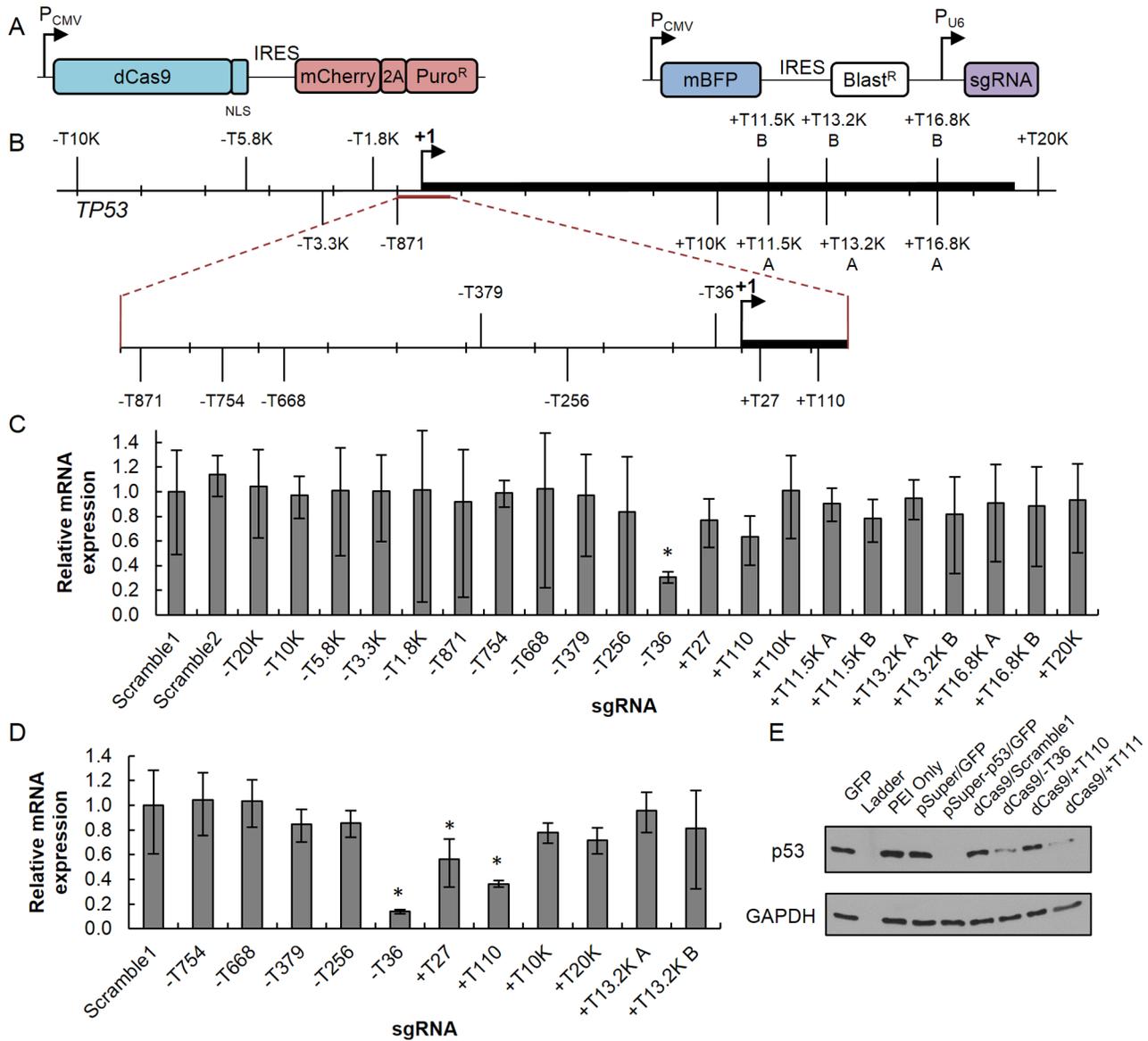


Figure 1. Repression of *TP53* by targeting dCas9 to the transcriptional start site. (A) HEK 293T cells were co-transfected with dCas9 (left) and sgRNA (right) expression plasmids. Codon-optimized dCas9 was fused to three copies of nuclear localization signal (NLS) and was co-expressed with mCherry fluorescent protein. sgRNA plasmid expresses mBFP and sgRNA off separate promoters. P_{CMV}, CMV promoter; 2A, ribosomal slippage site; Puro^R, puromycin resistance gene; IRES, internal ribosome entry site; mBFP, TagBFP fluorescent protein; Blast^R, blasticidin resistance gene; P_{U6}, U6 promoter. (B) Locations of sgRNA binding sites in the *TP53* promoter and transcribed region (thick line). Each sgRNA is numbered by the distance (bp) from the transcriptional start site. “-”, upstream of +1; “+”, downstream of +1; “K”, one thousand bp; +1, transcriptional start site. Labeled sites above and below the transcribed region indicate sgRNAs targeting the template or non-template DNA strands, respectively. (C, D) Relative expression of *TP53* mRNA in cells co-transfected with dCas9 and indicated sgRNA constructs. After three days, cells were (C) directly analyzed by qRT-PCR or (D) sorted for BFP-positive cells, then analyzed by qRT-PCR. Results were normalized, linearly rescaled, and calculated for mean fold change ($n = 3$) \pm 95% confidence interval, relative to Scramble1 negative control sgRNA. * $P < 0.01$ compared to non-targeted sgRNA control by paired, one-sided t-test. (E) Immunoblot from HEK 293T cell lysate three days post-transfection with shRNA against *TP53*, dCas9 and sgRNAs against *TP53* (e.g., dCas9/-T36), dCas9 and non-targeted sgRNA (dCas9/Scramble1), other control constructs (GFP expression vector, which served as non-specific vector control, or pSuper without any shRNA encoded), or mock (PEI only). See Figure 2E for location of sgRNA +T111. pSuper constructs were co-transfected with GFP-IRES-mCherry-2A-Puro expression constructs for selection purposes and experimental consistency (pSuper/GFP, pSuper-p53/GFP). Protein ladder (lane 2 from left) is not visible. See Figure S5 for uncropped immunoblots. doi:10.1371/journal.pone.0113232.g001

To summarize, our results demonstrate significant CRISPRi repression of *TP53* mRNA levels without use of a KRAB effector domain by targeting the TSS. Modest repression was also observed in four of five genes targeted at the TSS. Our study demonstrates that target site efficiencies can vary greatly. When one desires to utilize CRISPRi to repress genes in mammalian

cells, one should evaluate a range of sgRNA target sites, including those near or at the transcriptional start site. However, the targeting of the transcriptional start site is not guaranteed to be effective for every gene and thus will need to be evaluated on a gene-by-gene basis. We propose that CRISPRi with dCas9 alone may prove useful for functional mapping of transcription factor

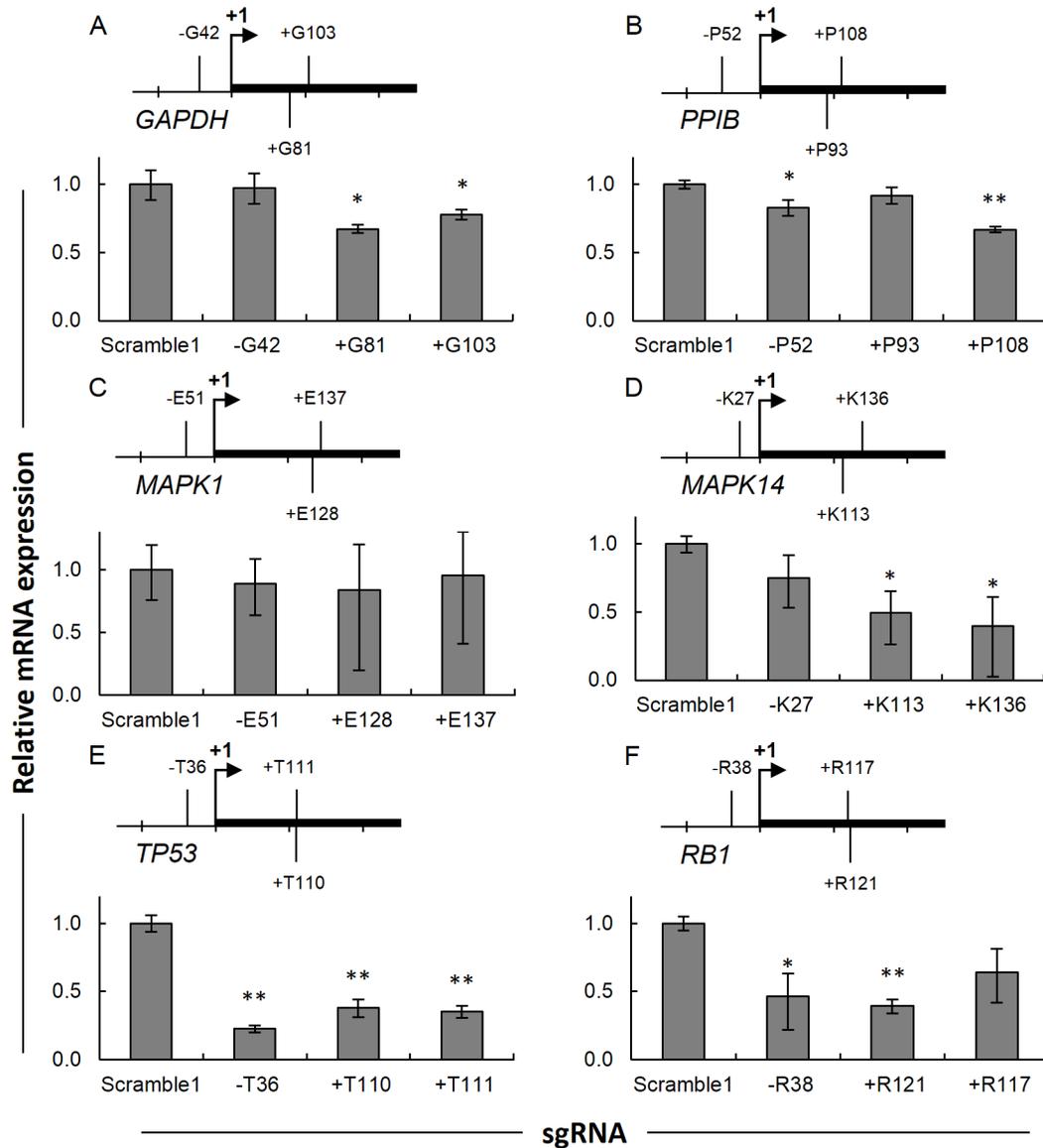


Figure 2. Evaluation of dCas9 targeting to the TSS of various genes. Diagram of sgRNA binding sites and plots of mRNA repression in HEK 293T by dCas9 and sgRNAs targeting (A) *GAPDH*, (B) *PPIB*, (C) *MAPK1*, (D) *MAPK14*, (E) *TP53*, and (F) *RB1*. Each sgRNA is numbered by the distance from the transcriptional start site. “-”, upstream of +1; “+”, downstream of +1; +1, transcriptional start site. Labeled sites above and below the transcribed region indicate sgRNAs targeting the template or non-template DNA strands, respectively. Relative expression of each mRNA in HEK 293T cells co-transfected with dCas9 and sgRNA constructs. After three days, cells transfected with dCas9 and indicated sgRNA constructs were sorted for GFP-positive cells, analyzed by qRT-PCR. Results were normalized, linearly rescaled, and calculated for mean fold change ($n = 3$) \pm 95% confidence interval, relative to Scramble1 negative control sgRNA. * $P < 0.02$, ** $P < 0.002$ compared to Scramble1 sgRNA control by paired, one-sided t-test. doi:10.1371/journal.pone.0113232.g002

motifs and enhancer elements via transcription initiation blocking without the potential additional repressive effect of a KRAB domain.

Supporting Information

Figure S1 Evaluation of CRISPR-mediated repression using dCas9 and dCas9 fused to a KRAB repressor domain. (A) HEK 293T cells were co-transfected with dCas9 or dCas9 fused to KRAB domain (left) and sgRNA (right) expression plasmids. Both codon-optimized dCas9 and dCas9-KRAB were fused to three copies of nuclear localization signal (NLS) and were co-expressed with mCherry fluorescent protein. sgRNA plasmid

expresses mBFP and sgRNA off separate promoters. P_{CMV} , CMV promoter; 2A, ribosomal slippage site; Puro^R, puromycin resistance gene; IRES, internal ribosome entry site; mBFP, TagBFP fluorescent protein; Blast^R, blasticidin resistance gene; P_{U6} , U6 promoter. (B) Locations of sgRNA binding sites in the endogenous *TP53* locus. Each sgRNA is numbered by the distance (bp) from the transcriptional start site. “-”, upstream of +1; “+”, downstream of +1; “K”, one thousand bp; +1, transcriptional start site. Labeled sites above and below the transcribed region indicate sgRNAs targeting the template or non-template DNA strands, respectively. (C–E) Relative expression of *TP53* mRNA in cells co-transfected with dCas9 and sgRNA constructs targeting (C) upstream, (D) downstream of the +1 site or

(E) in combinations of multiple sgRNA. After three days, cells co-transfected with indicated dCas9 and sgRNA constructs were analyzed by qRT-PCR. Data in (C, D) are fold change relative to Scramble1 or Scramble2 negative control sgRNA \pm s.e. of three technical replicates. Data in (E) represents sorted cells and were normalized, linearly rescaled, and calculated for mean fold change ($n = 3$) $\pm 95\%$ confidence interval, relative to Scramble1 negative control sgRNA. * $P < 0.01$ compared to non-targeted sgRNA control by paired, one-sided t-test. See also Figure 1B for additional targeted *TP53* sites.

(TIF)

Figure S2 CRISPR-mediated repression of GFP. Destabilized GFP (ds1GFP) expression cassette expressed using an (A) LTR (P_{LTR}) or (B) mutant EF-1 α promoter (P_{EF-T05}) was retrovirally transduced into HEK 293 cells and selected with neomycin. LTR, retroviral long-terminal repeat; Neo^R, neomycin resistance gene; IRES, internal ribosome entry site. Cells were analyzed three days post-co-transfection with dCas9-KRAB and either GFP1 or GFP2 sgRNAs targeting the template strand. GFP fluorescence was measured via flow cytometry after gating for GFP positive cells. Values are arithmetic means of GFP fluorescence \pm s.d. ($n = 3$) calculated from geometric means of each sample population and were normalized to dCas9:Scramble1 negative control.

(TIF)

Figure S3 Target sequence -T36 also leads to a reduction in *WRAP53 α* isoform mRNA. (A) Locations of sgRNA binding sites in the endogenous *TP53* locus. Each sgRNA is numbered by the distance (bp) from the transcriptional start site. “-”, upstream of +1; “+”, downstream of +1; “K”, one thousand bp; +1, transcriptional start site. Labeled sites above and below the transcribed region indicate sgRNAs targeting the template or non-template DNA strands, respectively. (B, C) Relative expression of *TP53*, *WRAP53 α* , and all isoforms of *WRAP53* mRNA in cells co-transfected with (B) dCas9 and sgRNA constructs or (C) shRNA constructs targeting *TP53*. After three days, cells co-transfected with indicated constructs were sorted, analyzed by qRT-PCR. Data are fold change relative to (B) Scramble1 negative control sgRNA or (C) pSuper Control plasmid \pm s.e. of three technical replicates.

(TIF)

Figure S4 Annealing temperature of sgRNA target sequence has minimal to no correlation with reduction in transcriptional expression. Minimal to no correlation was

observed between transcriptional repression relative to Scramble1 control of all individual dCas9 knockdown experiments done in triplicate and the annealing temperature of the sgRNA target. Line represents linear regression of data.

(TIF)

Figure S5 Reduction of p53 protein in transfected HEK 293T cells. Uncropped immunoblot (see Figure 1E) containing 15 μ g total protein/lane immunostained with (A) p53 antibody and (B) GAPDH antibody from HEK 293T cells transfected as indicated. Protein ladder (lane 2 from left) is not visible.

(TIF)

Table S1 Target sequences for sites within and flanking human *TP53*. The leading G nucleotide required for U6 promoter expression is in bold. The underlined following 19 to 24 nucleotides comprise the target sequence. The PAM site is in gray.

(TIF)

Table S2 Control target sequences. The leading G nucleotide required for U6 promoter expression is in bold. The underlined following 19 to 24 nucleotides comprise the target sequence. The PAM site is in gray.

(TIF)

Table S3 Target sequences within the promoter of various human genes of interest. The leading G nucleotide required for U6 promoter expression is in bold. The underlined following 19 to 24 nucleotides comprise the target sequence. The PAM site is in gray.

(TIF)

Table S4 Primer sequences used for quantitative RT-PCR.

(TIF)

Acknowledgments

Special thanks are owed to Jane Yang for helpful discussion and mentoring. We also thank the staff at the Stanford Shared FACS Facility for their help with cell sorting.

Author Contributions

Conceived and designed the experiments: IEBL JPF CLW. Performed the experiments: IEBL. Analyzed the data: IEBL. Contributed reagents/materials/analysis tools: IEBL JPF. Wrote the paper: IEBL JPF CLW.

References

- Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, et al. (2007) CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science* 315: 1709–1712. doi:10.1126/science.1138140.
- Wiedenheft B, Sternberg SH, Doudna JA (2012) RNA-guided genetic silencing systems in bacteria and archaea. *Nature* 482: 331–338. doi:10.1038/nature10886.
- Marraffini LA, Sontheimer EJ (2010) CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat Rev Genet* 11: 181–190. doi:10.1038/nrg2749.
- Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, et al. (2011) CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 471: 602–607. doi:10.1038/nature09886.
- Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, et al. (2012) A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* 337: 816–821. doi:10.1126/science.1225829.
- Mojica FJM, Díez-Villaseñor C, García-Martínez J, Almendros C (2009) Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* 155: 733–740. doi:10.1099/mic.0.023960-0.
- Shah SA, Erdmann S, Mojica FJM, Garrett RA (2013) Protospacer recognition motifs: Mixed identities and functional diversity. *RNA Biol* 10: 891–899. doi:10.4161/rna.23764.
- Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA (2013) RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol* 31: 233–239. doi:10.1038/nbt.2508.
- Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, et al. (2013) One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR/Cas-Mediated Genome Engineering. *Cell* 153: 910–918. doi:10.1016/j.cell.2013.04.025.
- Mali P, Yang L, Esvelt KM, Aach J, Guell M, et al. (2013) RNA-Guided Human Genome Engineering via Cas9. *Science* 339: 823–826. doi:10.1126/science.1232033.
- Jinek M, East A, Cheng A, Lin S, Ma E, et al. (2013) RNA-programmed genome editing in human cells. *eLife* 2: e00471. doi:10.7554/eLife.00471.
- Cong L, Ran FA, Cox D, Lin S, Barretto R, et al. (2013) Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* 339: 819–823. doi:10.1126/science.1231143.
- Hwang WY, Fu Y, Reyon D, Maeder ML, Tsai SQ, et al. (2013) Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat Biotechnol* 31: 227–229. doi:10.1038/nbt.2501.
- Xu K, Ren C, Liu Z, Zhang T, Zhang T, et al. (2014) Efficient genome engineering in eukaryotes using Cas9 from *Streptococcus thermophilus*. *Cell*

- Mol Life Sci. Available: <http://link.springer.com/10.1007/s00018-014-1679-z>. Accessed 2014 August 18.
15. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, et al. (2013) Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell* 152: 1173–1183. doi:10.1016/j.cell.2013.02.022.
 16. Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, et al. (2013) CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell* 154: 442–451. doi:10.1016/j.cell.2013.06.044.
 17. Bikard D, Jiang W, Samai P, Hochschild A, Zhang F, et al. (2013) Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res* 41: 7429–7437. doi:10.1093/nar/gkt520.
 18. Kearns NA, Genga RMJ, Enuameh MS, Garber M, Wolfe SA, et al. (2014) Cas9 effector-mediated regulation of transcription and differentiation in human pluripotent stem cells. *Development* 141: 219–223. doi:10.1242/dev.103341.
 19. Maeder ML, Linder SJ, Cascio VM, Fu Y, Ho QH, et al. (2013) CRISPR RNA-guided activation of endogenous human genes. *Nat Methods* 10: 977–979. doi:10.1038/nmeth.2598.
 20. Cheng AW, Wang H, Yang H, Shi L, Katz Y, et al. (2013) Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. *Cell Res* 23: 1163–1171. doi:10.1038/cr.2013.122.
 21. Perez-Pinera P, Kocak DD, Vockley CM, Adler AF, Kabadi AM, et al. (2013) RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat Methods* 10: 973–976. doi:10.1038/nmeth.2600.
 22. Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, et al. (2013) Dynamic Imaging of Genomic Loci in Living Human Cells by an Optimized CRISPR/Cas System. *Cell* 155: 1479–1491. doi:10.1016/j.cell.2013.12.001.
 23. Larson MH, Gilbert LA, Wang X, Lim WA, Weissman JS, et al. (2013) CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nat Protoc* 8: 2180–2196. doi:10.1038/nprot.2013.132.
 24. Farzadfard F, Perli SD, Lu TK (2013) Tunable and Multifunctional Eukaryotic Transcription Factors Based on CRISPR/Cas. *ACS Synth Biol* 2: 604–613. doi:10.1021/sb4000081r.
 25. Biegging KT, Mello SS, Attardi LD (2014) Unravelling mechanisms of p53-mediated tumour suppression. *Nat Rev Cancer* 14: 359–370. doi:10.1038/nrc3711.
 26. Brummelkamp TR, Bernards R, Agami R (2002) A system for stable expression of short interfering RNAs in mammalian cells. *Science* 296: 550–553. doi:10.1126/science.1068999.
 27. Hao D-L, Liu C-M, Dong W-J, Gong H, Wu X-S, et al. (2005) Knockdown of Human p53 Gene Expression in 293-T Cells by Retroviral Vector-mediated Short Hairpin RNA. *Acta Biochim Biophys Sin* 37: 779–783. doi:10.1111/j.1745-7270.2005.00107.x.
 28. Suzuki Y, Yamashita R, Nakai K, Sugano S (2002) DBTSS: DataBase of human Transcriptional Start Sites and full-length cDNAs. *Nucleic Acids Res* 30: 328–331. doi:10.1093/nar/30.1.328.
 29. Yamashita R, Sugano S, Suzuki Y, Nakai K (2012) DBTSS: DataBase of Transcriptional Start Sites progress report in 2012. *Nucleic Acids Res* 40: D150–D154. doi:10.1093/nar/gkr1005.
 30. Tsuchihara K, Suzuki Y, Wakaguri H, Irie T, Tanimoto K, et al. (2009) Massive transcriptional start site analysis of human genes in hypoxia cells. *Nucleic Acids Res* 37: 2249–2263. doi:10.1093/nar/gkp066.
 31. Ferreira JP, Overton KW, Wang CL (2013) Tuning gene expression with synthetic upstream open reading frames. *Proc Natl Acad Sci* 110: 11284–11289. doi:10.1073/pnas.1305590110.
 32. Graham FL, Smiley J, Russell WC, Nairn R (1977) Characteristics of a Human Cell Line Transformed by DNA from Human Adenovirus Type 5. *J Gen Virol* 36: 59–72. doi:10.1099/0022-1317-36-1-59.
 33. Stewart N, Bacchetti S (1991) Expression of SV40 large T antigen, but not small t antigen, is required for the induction of chromosomal aberrations in transformed human cells. *Virology* 180: 49–57.
 34. Ferreira JP, Peacock RWS, Lawhorn IEB, Wang CL (2011) Modulating ectopic gene expression levels by using retroviral vectors equipped with synthetic promoters. *Syst Synth Biol* 5: 131–138. doi:10.1007/s11693-011-9089-0.
 35. Willems E, Leyns L, Vandesompele J (2008) Standardization of real-time PCR gene expression data from independent biological replicates. *Anal Biochem* 379: 127–129. doi:10.1016/j.ab.2008.04.036.
 36. Groner AC, Meylan S, Ciuffi A, Zangger N, Ambrosini G, et al. (2010) KRAB-zinc finger proteins and KAP1 can mediate long-range transcriptional repression through heterochromatin spreading. *PLoS Genet* 6: e1000869. doi:10.1371/journal.pgen.1000869.
 37. Hagmann M, Georgiev O, Schaffner W (1997) The VP16 paradox: herpes simplex virus VP16 contains a long-range activation domain but within the natural multiprotein complex activates only from promoter-proximal positions. *J Virol* 71: 5952–5962.
 38. Mahmoudi S, Henriksson S, Corcoran M, Méndez-Vidal C, Wiman KG, et al. (2009) Wrap53, a Natural p53 Antisense Transcript Required for p53 Induction upon DNA Damage. *Mol Cell* 33: 462–471. doi:10.1016/j.molcel.2009.01.028.
 39. Polson A, Reisman D (2014) The bidirectional p53–Wrap53 promoter is controlled by common cis- and trans-regulatory elements. *Gene* 538: 138–149. doi:10.1016/j.gene.2013.12.046.
 40. Saldaña-Meyer R, Recillas-Targa F (2011) Transcriptional and epigenetic regulation of the p53 tumor suppressor gene. *Epigenetics* 6: 1068–1077. doi:10.4161/epi.6.9.16683.
 41. Selby CP, Drapkin R, Reinberg D, Sancar A (1997) RNA polymerase II stalled at a thymine dimer: footprint and effect on excision repair. *Nucleic Acids Res* 25: 787–793. doi:10.1093/nar/25.4.787.
 42. Gao X, Tsang JCH, Gaba F, Wu D, Lu L, et al. (2014) Comparison of TALE designer transcription factors and the CRISPR/dCas9 in regulation of gene expression by targeting enhancers. *Nucleic Acids Res*: gku836. doi:10.1093/nar/gku836.
 43. Wu X, Kriz AJ, Sharp PA (2014) Target specificity of the CRISPR-Cas9 system. *Quant Biol*. Available: <http://link.springer.com/10.1007/s40484-014-0030-x>. Accessed 2014 October 9.
 44. Wu X, Scott DA, Kriz AJ, Chiu AC, Hsu PD, et al. (2014) Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat Biotechnol* 32: 670–676. doi:10.1038/nbt.2889.
 45. Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, et al. (2013) High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol* 31: 822–826. doi:10.1038/nbt.2623.
 46. Pattanayak V, Lin S, Guilinger JP, Ma E, Doudna JA, et al. (2013) High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotechnol* 31: 839–843. doi:10.1038/nbt.2673.
 47. Cho SW, Kim S, Kim Y, Kweon J, Kim HS, et al. (2014) Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res* 24: 132–141. doi:10.1101/gr.162339.113.
 48. Kucsu C, Arslan S, Singh R, Thorpe J, Adli M (2014) Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat Biotechnol* 32: 677–683. doi:10.1038/nbt.2916.